NOAA Technical Report NOS CO-OPS 0020

NOS Procedures for Developing and Implementing Operational Nowcast and Forecast Systems for PORTS

Silver Spring, Maryland
January 1999

noaa National Oceanic And Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
National Ocean Service
Center for Operational Oceanographic Products and Services

Center for Operational Oceanographic Products and Services
National Ocean Service
National Oceanic and Atmospheric Administration
U.S. Department of Commerce

The National Ocean Service (NOS) Center for Operational Oceanographic Products and Services (CO-OPS) collects and distributes observations and predictions of water levels and currents to ensure safe, efficient and environmentally sound maritime commerce. The Center provides the set of water level and coastal current products required to support NOS' Strategic Plan mission requirements, and to assist in providing operational oceanographic data/products required by NOAA's other Strategic Plan themes. For example, CO-OPS provides data and products required by the National Weather Service to meet its flood and tsunami warning responsibilities. The Center manages the National Water Level Observation Network (NWLON), and a national network of Physical Oceanographic Real-Time Systems (PORTS) in major U.S. harbors. The Center: establishes standards for the collection and processing of water level and current data; collects and documents user requirements which serve as the foundation for all resulting program activities; designs new and/or improved oceanographic observing systems; designs software to improve CO-OPS' data processing capabilities; maintains and operates oceanographic observing systems; performs operational data analysis/quality control; and produces/disseminates oceanographic products.

NOAA Technical Report NOS CO-OPS 0020

NOS Procedures for Developing and Implementing Operational Nowcast and Forecast Systems for PORTS

January 1999

noaa National Oceanic and Atmospheric Administration

U.S. DEPARTMENT OF COMMERCE
William M. Daley, Secretary

National Oceanic and Atmospheric Administration
D. James Baker, Under Secretary

National Ocean Service
Nancy Foster, Assistant Administrator

Center For Operational Oceanographic Products and Services
Richard Barazotto, Chief

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

TABLE OF CONTENTS

INTRODUCTION

In order to meet its operational oceanographic mission responsibilities, the National Ocean Service/Center for Operational Oceanographic Products and Services (NOS/CO-OPS) will be developing and implementing nowcast and forecast models to support the Physical Oceanographic Real Time Systems (PORTS). These models will be designed to enhance the navigational guidance supplied by PORTS real-time observations by providing information regarding both the present (nowcast) and future (forecast) oceanographic conditions at many locations within an estuary. These models will be developed within CO-OPS, the Coast Survey Development Laboratory (CSDL) and by groups outside of NOS. NOS must ensure that these models have been developed and implemented in a scientifically sound and operationally robust way; that the model's shortcoming are understood; that the products are clear, understandable, and useful; and that all products and procedures are authoritative in the face of potential legal challenges. It is imperative that the nowcast and forecast systems are developed consistent with user needs and with the operational environment in which they will be run. All models (including statistical models) that produce NOS PORTS-sanctioned nowcasts and forecasts in support of safe navigation, whether developed within or outside NOS, will be developed and implemented in adherence to the procedures contained in this document.

Definition of a Nowcast/Forecast System

The following definitions are used for this document. A *hindcast, nowcast, and forecast* are scientific predictions about the past, present, and future states, respectively, of water levels and/or currents (and possibly other relevant oceanographic variables such as salinity and temperature) in a bay or estuary made by a numerical, statistical, or hybrid model or method. These predictions rely on either observed or forecast data, not on hypothetical data. A *hindcast* incorporates past or historical observational data. A *nowcast* incorporates recent (and often near real-time) observed meteorological, oceanographic, and/or river flow rate data; covers the period of time from the recent past (up to a few days) to the present; and makes predictions for locations where observational data are not available. The present is the time at which the nowcast is made, and at which the most recent observations are from a few minutes to an hour old. A *forecast* incorporates meteorological, oceanographic, and/or river flow rate forecasts; makes predictions for locations where observational data will not be available; and is usually initialized by the results of a nowcast.

Approach

These procedures are designed to guide the development stage and the implementation processes which consist of four broad phases. Several of the phases contain multiple stages and steps which will be described in the following sections of this document. The phases are:

! Problem Definition
! Development of a Nowcast/Forecast System
! Operation of a Nowcast/Forecast System
! Enhancement of a Nowcast/Forecast System

The overall philosophy of this process is iterative. It is designed to move from broad to specific requirements by utilizing user input at many stages of the process. Thorough review and documentation of each step is deemed critical (documentation requirements are covered in Appendix A). Model evaluation is an important part of development and is designed to determine how accurate the method is in reproducing historical observations and in predicting future behavior. Model evaluation and skill assessment methodology is discussed in detail in Appendix B.

Teams and Review Panels

Developing and implementing nowcast/forecast system will require the formation of several teams and review panels. The functions and composition of these teams are described below.

Steering Committee - A PORTS Forecast System Steering Committee (PFSSC) will be established to ensure adherence to the procedures contained in this document. The Committee will be composed of five members and will be chaired by the NOS Chief Scientist and will include the Directors of CO-OPS, Office of Coast Survey (CS), Office of Response and Restoration (OR&R), and the National Geodetic Survey. The Committee will oversee the development and implementation process. It will approve the following: membership of review panels, proposed modeling approaches, model accuracy/performance standards, testing procedures, and operational implementation plans. It will establish and modify practices as needed. It is expected that, as NOS gains nowcasting and forecasting experience, specific requirements described in this document, especially in the model skill assessment, will be modified or expanded to reflect knowledge gained.

System Design Team- The PORTS Development Manager (PDM) will assemble a System Design Team (SDT) with the skills required to address the problem at hand. The PFSSC will approve the membership of the SDT. At a minimum, this Team will include the Developer (who is the principle scientist), and participants from CO-OPS's Requirements and Development Division RDD and Information Systems Division (ISD), as well as representatives from the affected User Groups. The SDT may also include personnel from the Coast Survey Development Laboratory (CSDL), the National Centers for Coastal Ocean Science (NCCOS), OR&R, as well as non-NOS researchers. This Team is responsible for the design and development of the nowcast/forecast system.

System Implementation Team - The PDM will assemble and the Steering Committee will approve a System Implementation Team (SIT). This Team consists of the Developer, selected SDT members, and the CO-OPS personnel who will oversee the implementation of the system in an operational environment. This Team is responsible for the implementation of the forecast system in operational mode.

Review Panel - This Panel would be selected and convened by the PFSSC. Members of this Panel will include representatives from CO-OPS, NCCOS, OR&R, and CS/CSDL. It may also include representatives from the NWS National Centers for Environmental Prediction (NCEP), NWS Techniques Development Laboratory (TDL), as well as non-NOAA representatives from the User Group, U.S. Coast Guard, local universities, the U.S. Geological Survey, the U.S. Army Corps of Engineers, and the U.S. Navy. Reviewers could be contacted via Internet, conference calls, or other means. The Panel will review plans, documents, and products and make written comments and recommendations to the PFSSC.

User Group - This group consists of interested members of the local marine community who will directly use the products being developed. Partners who have been involved in the development and implementation of PORTS in the target estuary will be asked to participate in the User Group. It may also include pilots, port facility operators, U.S. Navy and U.S. Coast Guard personnel, NWS marine forecasters, marine resource managers, and recreational boaters.

Beta Testers - The Beta Testers are selected by the SDT and approved by the PFSSC and are a subset of the User Group. The Testers will evaluate the experimental product, keep records, and make an impartial analysis. The Testers will make written comments and recommendations to the SDT and PFSSC.

Operational Evaluators - The Operational Evaluators are selected by the SIT and approved by the PFSSC and are a subset of the User Group consisting of individuals representing the user organizations. The Operational Evaluators will evaluate the operational product on a frequent basis, keep records, and make an impartial analysis. They will make written comments and recommendations to the SDT and PFSSC.

Resulting Documents

The following official documents will be generated during the development and implementation of a nowcast/forecast system:

- Needs Assessment
- Requirements Document
- System Development Plan
- System Documentation
- Operational System Implementation Plan
- Operations and Maintenance Manual
- User's Guide

These written records will be described in the following sections and a listing of their expected contents is contained in Appendix A. The Requirements Document, the System Development Plan, and the System Description are *evolving* documents since contributions to them will be made throughout the development phase. The Operational System Implementation Plan and the Operations and Maintenance Manual are *static* documents written at the completion of a specific stage. In addition, formal Reviewer Comments will be written up for the record (Appendix A).

Application to Existing Nowcast/Forecast Systems

For model development in support of PORTS which is already underway in NOS and in other institutions, the specific modelling approach will have already been selected. However, the Requirements Document and the System Description will still be required. A SDT will be assembled to produce these documents. The PFSSC will review the state of development of the existing system. The Committee will decide which procedures have been successfully completed, and which procedures need to be completed before the system gains NOS acceptance.

PHASE 1: PROBLEM DEFINITION

Many factors will be considered to make the determination of whether a given PORTS site is a candidate for model development, including the spatial variability of the circulation and the number of requests for information that are not satisfied by real time instrumentation alone. Once a PORTS site has been designated as a potential candidate for the addition of nowcast/forecast capabilities, the PDM will initiate the Problem Definition Phase. The PDM will define and document user requests and requirements. Requirements may refer to written requests for products or verbal requests from specific user groups. The PDM will analyze the needs, review potential responses, and assess alternatives approaches.

Following the PDM's analysis, the Director CO-OPS will formulate recommendations for approval by the PSSFC. The PDM will draft a preliminary Needs Assessment for the particular port. The PFSSC will review/approve the document and forward it, with recommendations on a course of action (options include no further action, referring the requesters to other providers, or NOS involvement), to the Director, CO-OPS. If CO-OPS decides that continued development is warranted, the PDM will complete a final Needs Assessment and the process moves to the System Development Phase.

PHASE 2: DEVELOPMENT OF A NOWCAST/FORECAST SYSTEM

The System Development Phase is composed of four stages: analysis of users' functional requirements and needs, selection of candidate solution, development of demonstration nowcast/forecast system, and development of prototype nowcast/forecast system. After being tested and reviewed, the prototype system becomes the basis for the operational system.

Stage 1: Analysis of Requirements

Following PFSSC approval of the recommendations of the Needs Assessment document, the PDM will then assemble a System Design Team (SDT) to respond to users' nowcast and forecast requirements and needs, discuss the scientific issues, and examine operational scenarios of a potential NOS forecast system. Using the Needs Assessment as a base, the SDT's first task will be to draft a *preliminary* version of a Requirements Document. This will likely require contacting key users to gain more information. This document will be provided to the PFSSC for approval.

After approval, the SDT will present a summary of this document at the *first joint NOS/Users meeting* at the site of the port. The Team will show examples of NOS forecast systems at other ports to illustrate NOS capabilities. This meeting serves as a 'reality check' to ensure that the SDT has clearly defined the users forecast information requirements for their particular port and to gain initial user support for the developing a NOS forecast system.

The SDT will incorporate Users' comments from the meeting into the *final* version of the Requirements Document. This document will be given to the PFSSC for review and approval.

Stage 2: Selection of a Candidate Solution

Once approved, the Requirements Document will be used by the SDT to assess approaches. Potential approaches include statistical model-based, numerical hydrodynamic model-based, and artificial intelligence-based (i.e. expert systems and neutral networks) forecast systems. The Team will then select a specific approach or candidate solution to meet the forecast information needs of the users.

Once the approach is selected, the SDT will generate a *broad specification* of the nowcast/forecast system requirements. System requirements would include the following: (1) model requirements (e.g., if a numerical model, the horizontal and vertical grid resolution, surface and open boundary forcing method); (2) real-time meteorological and oceanographic observations and forecast requirements; (3) dissemination requirements (e.g., methods and timeliness); and (4) computer, network, and internal/external personnel requirements.

The SDT, with the Developer as the lead, will draft a *preliminary* version of a System Development Plan which will include descriptions of these nowcast/forecast system requirements, a plan to meet them, and a work schedule. The SDT will present the Plan to the PFSSC, which will review it. If the PFSSC accepts the Plan, the SDT will move on to the development of a Demonstration System for the port.

Stage 3: Construction of the Demonstration System

The purpose of the Demonstration System is to demonstrate the feasibility and accuracy of the approach. The Demonstration System allows both the SDT and User Group to have a better understanding of the users' requirements, the different strengths and weaknesses of different modeling methods, product design, and operational constraints before major investments are made. The Demonstration System will also help transfer the broad specification into a tangible system that can then be refined to obtain a *detailed specification* of the system. The Demonstration System relies solely on archived data for test nowcasts (i.e, hindcasts) and test forecasts.

The SDT will use the *preliminary* System Development Plan (developed in Stage 2) to build a Demonstration System. The Demonstration System will be constructed and calibrated, then run in hindcast mode for a specified time period and its test nowcasts and forecasts evaluated (see Appendix B).

Requirements for additional sensors may be generated during the development of a Demonstration System. In this event, the SDT will prepare a brief description and justification of the additional instrumentation (either temporary sensors deployed for model verification and/or PORTS real time sensors) for the consideration by the CO-OPS PORTS Management Team. After the additional instrumentation is deemed warranted, the recommendations will be passed to the PORTS Development Team which will use this information to work with the local sponsors of PORTS and with CO-OPS to implement the enhancements.

The SDT will prepare a *preliminary* version of the System Documentation. This will include information regarding the prediction equations, inputs and outputs, samples products, forecast skill, and product dissemination requirements. After approval of the document by the PFSSC, the SDT will present their Demonstration System to the first joint meeting with PFSSC and the Review Panel, along with copies of their *preliminary* System Documentation and System Development Plan. Since it unlikely that the user representatives or external developers on the SDT will be able to attend all meetings of the Team, a *restricted* WWW site will be established in order to easily display forecast system output. Comments from the user representatives and external developers will be obtained using the WWW site in conjunction with telephone conference calls.

When the Demonstration System and its documentation are found to be acceptable by the PFSSC, the SDT will present the system including sample products to a *second joint meeting of NOS and Users* at the site of the port. The purpose of this meeting to develop a *detailed specification* of the forecast system based on user comments and to also maintain user support for the forecast system. The SDT will also be encouraged to present their system at their professional society's conferences/workshops and in journals in order to obtain feedback from peers outside of NOS.

The SDT will draft a *final* version of the System Development Plan that will include a detailed specification of the forecast system. The final version will incorporate the users' comments obtained from the second meeting, comments from the Review Panel along with greater emphasis on operational issues. The SDT will present their Plan to the PFSSC and the Review Panel. If PFSSC approves of the Plan, the Team would move on to the development of a prototype forecast system for the port.

Stage 4: Construction of the Prototype System

The purpose of the Prototype System is to refine the system to meet the detailed specification, field test the system, and make it the basis for an operational system. Thus, the Prototype System should closely approximate the operational environment in order to minimize the time required to implement the system operationally. The Prototype System relies primarily on the latest available forecasts of winds and water levels for its forecasts.

The SDT will use the *final* System Development Plan to build the Prototype Forecast System. The experimental system will be run in a semi-operational mode (e.g., 5 day/week) consisting of semi-operational nowcast and forecast cycles. It will be run in this manner for at least 3 months and its reliability, skill, and products evaluated. The products will be display on the restricted WWW site in order to obtain feedback from the Team's user representatives and external developers.

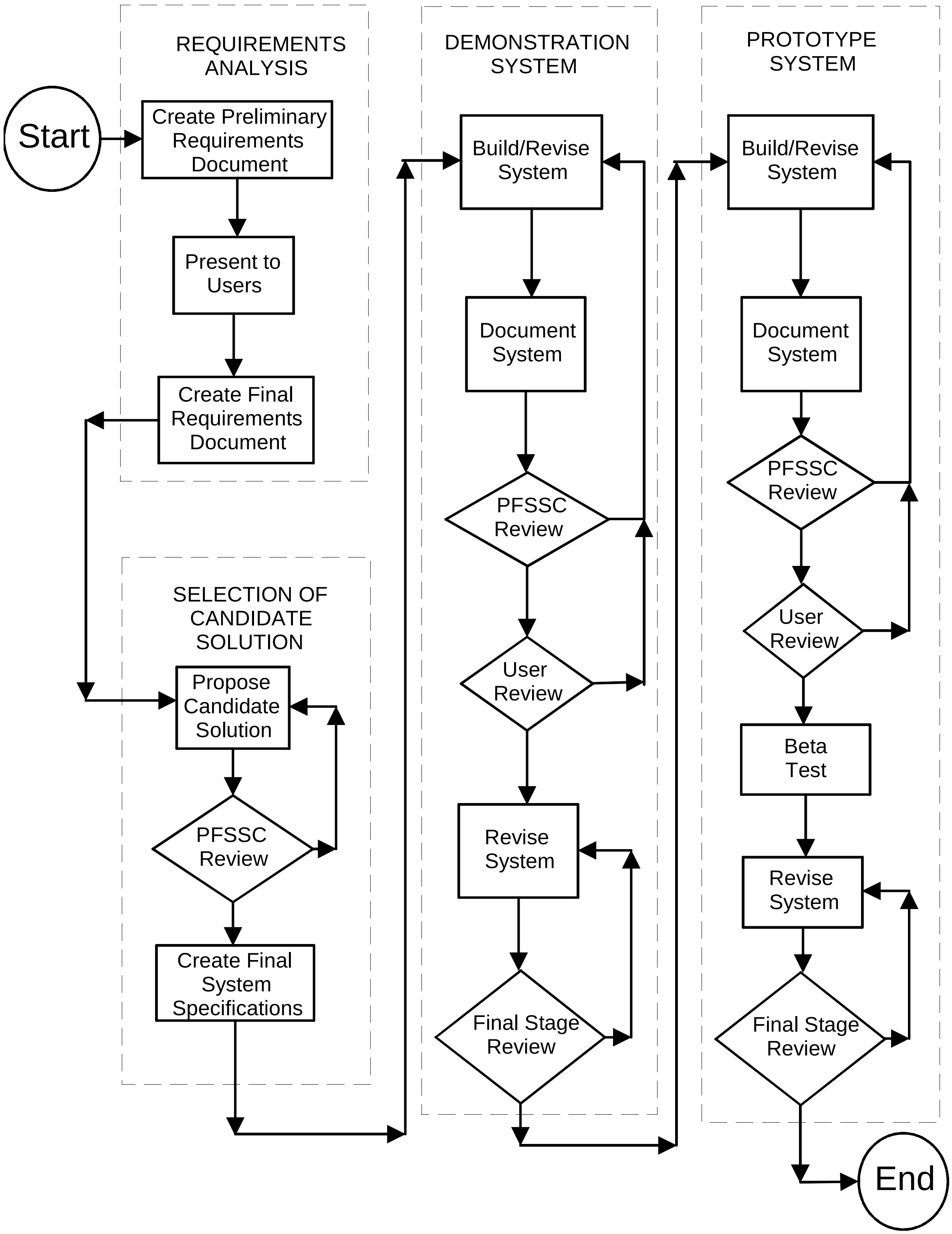
Once the SDT concludes that the system is producing accurate and reliable nowcasts and forecasts, the Team will update their Model Documentation and present the prototype system to a meeting with the PFSSC and the Review Panel. As in the previous stage, the SDT is encouraged to present their experimental system at their professional society's conferences or workshops and in journals in order to obtain feedback from peers outside of NOS.

If the PFSSC approves the system, the SDT will (1) present the system and its products at a *third joint NOS/User Group meeting* at the site of the port and (2) display the products on an *open* WWW site. The purpose of this meeting is to inform users of the availability of the semi-operational forecast products and to solicit and select volunteers to be Beta Testers for a specific time period (i.e. 6 months or less). The SDT will document the results of the beta test as an appendix in their System Documentation and use the results in making refinements to the experimental system.

Following the beta test, the SDT will prepare a final versions of the System Implementation Plan and System Documentation. The System Design Team will present the Prototype System to a meeting with the PFSSC and the Review Panel. The Panel will advise the PFSSC as to whether they consider the system is ready to move into the Operational Phase.

If PFSSC decides to move the forecast system into the Operational Phase, many of the System Design Team members will become members of the Implementation Team. In addition, the Beta Testers will be requested to continue as Operational Evaluators in the next phase. A flowchart of the steps in Phase 2 appears in Figure 1.

PHASE 2: SYSTEM DEVELOPMENT



PHASE 3: OPERATION OF A NOWCAST/FORECAST SYSTEM

Stage 1: Transition to Operations

The SDT will initiate the transition of the Prototype System to operational status by convening the System Implementation Team (SIT), which will orchestrate the physical implementation. The Implementation Plan for the transition to the operational environment will be developed by the SIT. The Plan will assign tasks and responsibilities for implementation and will identify the resources required. The PFSSC will review the Implementation Plan and discuss the required resources with the effected NOS organizations. Actual implementation will follow approval of the plan by the PFSSC.

Stage 2: Implementation

Using the strategy provided in the Implementation Plan, the SIT will migrate all system and application software, software interfaces, and hardware to the operational environment. The Team will implement the data management strategy that is defined in the Implementation Plan, including the provision for archiving of all system inputs and outputs. The Team will test the system and insure that all subsystems are operating as specified. Once certified, the hardware, software, and interfaces will become the Operational System.

The Developer, in cooperation with CO-OPS personnel, will produce an Operations and Maintenance Manual and a Users' Guide. The Users' Guide will be reviewed by the PFSSC and members of the User Group before open public distribution.

Stage 3: Evaluation of the Operational System

Evaluation of the Operational System will be conducted for a minimum of one month following the successful implementation and documentation of the system. During this 'shake-down' period the timeliness and reliability of inputs and outputs will be monitored by the Implementation Team. Changes in the implementation as a result of this operational evaluation will require retesting and subsequent evaluation. Once the system is deemed operationally ready by the PFSSC, the system will be officially commissioned and full public dissemination will commence.

Stage 4: Monitoring

Continuous monitoring of the data streams, forecast products, and dissemination will be the responsibility of CO-OPS. This function will be carried out by the Continuous Operational Real-Time Monitoring System (CORMS) (Gill et al., 1997). Nowcast and forecast inconsistencies, relative to criteria provided by the SDT, will be noted by CORMS and passed to the SDT for resolution. The provision to terminate any display of operational nowcast and/or forecast data or products from PORTS will be included in the monitoring implementation.

Stage 5: Data Management

All PORTS data and information products disseminated to the public will be stored in the National PORTS Database (NPDB) (Bethem, 1998a, b). During the design phase the SDT will meet with the ISB Database Administrator to insure that the new model data output conforms to NPDB standards. Some software will be provided to Developer to prepare the data for conformance to the PORTS data model and for bulk loading the database. The Developer will also be required to provide the necessary metadata used in some of the NPDB tables. Short- and long-term archival of the product will be managed by the database which will be the primary and only certified copy of the output product data.

In addition, the gridded model nowcast/forecasts fields will be packed in the World Meteorological Organization's Gridded Binary (GRIB) format (Dey, 1996). GRIB is the standard format used by all national meteorological and oceanographic centers. GRIB files will be disseminated using NOAA communication lines and will be available to NOAA's National Oceanographic Data Center for archiving.

PHASE 4: ENHANCEMENT OF A NOWCAST/FORECAST SYSTEM

The Users and the SDT will continue to assess system results and, if deemed necessary, make requests for system changes or enhancements to the PFSSC. Changes may be required when user requirements are updated, problem forecast situations are encountered, new forecast techniques are developed, updated model performance statistics become available, or new atmospheric or ocean models become available. All requests for major enhancements to an existing operational nowcast/forecast system will be evaluated by the PFSSC, which will decide the starting point in the procedures which is appropriate for the enhancement. Minor enhancements and corrections will be submitted directly to the SDT or the SIT, whichever appropriate. All requests will be in writing, and the response and/or changes to the system will be documented in the Operational System Implementation Plan and/or the Operations and Maintenance Manual.

ACKNOWLEDGMENTS

This report is a product of a cooperative effort between personnel from CO-OPS and CSDL. CSDL's Kurt Hess chaired the Team and was responsible for editing the drafts and for developing most of the Model Evaluation policy and procedures. Sonny Richardson (CO-OPS contractor) developed the initial approach to system development and implementation. CO-OPS's Kathryn Bosley and Tom Bethem contributed important ideas and concepts to the Procedures, as did CSDL's Frank Aikman III. CSDL's John Kelley tirelessly proofread the drafts and, more importantly, devised the concepts of demonstration and prototype systems. CSDL's Bruce Parker supplied additional guidance for the model evaluation process. CSDL's Lloyd Huff helped identify relevant variables and devise system accuracy requirements.

BIBLIOGRAPHY

Bethem, T.,1998a. Systems Development Plan: National PORTS Database. NOS, Oceanographic Products and Services Division. 20 pp.

Bethem, T., 1998b. Implementation Plan: National PORTS Database. NOS, Oceanographic Products and Services Division. 22 pp.

Bosley, K. T., T. Bethem, and M. J. Evans, 1997. System Development Plan. Chesapeake Area Forecasting Experiment (CAFE): Migration to an Operational Product. NOS, Oceanographic Products and Services Division, 26 pp + Appendix.

Dey, C.H., 1996. The WMO Format for the Storage of Weather Product Information and the Exchange of Weather Product Messages in Gridded Binary Form. National Centers for Environmental Prediction, Office Note 388, NCEP Central Operations, Camp Springs, MD.

Gill, S., W. Stoney, and T. Bethem, 1997. System Development Plan. CORMS: Continuous Operational Real-Time Monitoring System. NOAA Technical Report NOS OES 14. U.S Department of Commerce, NOAA. 41 pp.

Hicks, S., 1989. Tide and Current Glossary. National Ocean Service. 30 pp.

Kelley, J. G. W., E. J. Wei, and S. Maxwell, (in preparation). Operational Data Acquisition and Archive System (ODAAS). NOS, Coast Survey Development Laboratory.

Preisendorfer, R. W., 1988. Principal Component Analysis in Meteorology and Oceanography. Developments in Atmospheric Science, v. 17. Elsevier, Amsterdam. pp 425.

Zervas, C., (in preparation). Tidal Current Analysis Procedures and Associated Computer Programs. NOS, Coast Survey Development Laboratory.

APPENDIX A. DOCUMENTATION

Most of the steps in the forecast system development will be accompanied by a written document; these are referred to in the text in the relevant sections. The documents created during this process (in the order they are required) are listed below. In addition, upon completion of each formal review, participants will submit comments in writing to the PFSSC. Comments will summarize the (1) system being presented; (2) observations, likes and dislikes, and problems; and (3) recommendations.

Needs Assessment

The PORTS Development Manager (PDM), with the assistance of other NOS personnel as needed, will create a forecast Needs Assessment for the particular port. The purpose is to document and analyze user forecast information requirements and recommend a course of action. The document includes: (1) preliminary assessment of users' informational needs for the port, including copies of letters from users; (2) list of potential user/customer groups; (3) background information (e.g., current and previous NOS research and operational activities in the port); (4) the availability of real-time meteorological and oceanographic observations and forecasts in region of the port; and (5) a recommendation on a course of action.

Requirements Document

The System Design Team will create a Requirements Document, the purpose of which is to discuss user needs and to propose a system which will meet those needs. This will include (1) a detailed list of users' forecast requirements and needs, (2) a list of alternatives to meet users' informational needs; (3) estimated costs of alternatives; (4) a discussion of oceanographic and meteorological issues relevant for forecasting, (5) a proposed operational scenario of a potential forecast system, (6) examples of NOS forecast systems at other estuaries or ports to illustrate NOS capabilities, (7) users' comments from the 'reality check' meeting; (8) interested collaborators (i.e. researchers at local universities, colleges, or other federal/state agencies); (9) assessment of current resources in NOS for developing the forecast system; and (10) recommendations. The detailed user requirement section will include a list of predictands, forecast interval and length, locations, timeliness, and modes of dissemination.

System Development Plan

The System Design Team will lead the writing of the System Development Plan. The goal of this document is to define a rationale and plan for the development of a nowcast/forecast system to meet user needs. It should contain: (1) an introduction which describes the user requirements being addressed; (2) a description of the selected nowcast/forecast method; (3) a section detailing system requirements, including proposed dissemination methods; (4) the design of the proposed system; (5) a plan for the development of the system; (6) the resources required for development (human and fiscal); and (7) a schedule for development.

System Documentation

The Model Documentation will be written by the Developer with the assistance of the System Design Team. The purpose of this document is to describe the model and system so its suitability, reliability, and accuracy can be assessed. The document will include (1) a summary of requirements and an explanation of how the approach will satisfy them; (2) the approach or methodology, including oceanographic and meteorological background, model equations and grid, input data, typical outputs, and potential products; (3) skill assessment including variables used, data, results, and comparison with the accuracy of other forecast systems; and (4) operational considerations such as a nowcast/forecast cycle description, computer and communications needs, input data needs, products, and methods of dissemination.

Operational System Implementation Plan

It is the responsibility of the Implementation Team to lead the writing of this document, with significant contributions from NOS personnel who will be carrying out many of the implementation tasks. The goal of the Implementation Plan is to provide a blueprint for the successful transition of a nowcast/forecast system to operations. It assigns tasks and responsibilities for implementation. This document will contain: (1) an introduction which details the relationship of the system to the System Development Plan, (2) an implementation strategy, (3) a planning section which details the various teams that will be established and the critical decisions which must be made, (4) sections which outline the various tasks categories, (5) a plan for testing and evaluation of the system, (6) the resources required for implementation (human and fiscal), (7) the data management strategy, and (8) a schedule for system procurement and setup.

Operations and Maintenance Manual

The Implementation Team, in cooperation with additional NOS personnel, will produce an Operations and Maintenance Manual. This manual should include (1) an overview of the system, (2) a definition of input and output data streams, (3) timing requirements for data streams and model runs, (4) computer resource requirements, (5) contingencies for missing input data, and (6) established back-up procedures. This manual will conform to existing NOS standards in form and content and will be updated as needed, based on feedback from system operators.

Users' Guide

This document will be written by the System Design Team for the purpose of summarizing the approach and describing in detail the forecast product (with its limitations) for a broad audience. It will contain (1) a description of the problem and the requirements; (2) a summary of the approach, methods, definitions, and input data; (3) a description of the product being disseminated; (4) the strength and weaknesses of the nowcast/forecast and a summary of skill; (5) references; and (6) a method for providing user feedback (e.g., an e-mail address) on the product.

APPENDIX B. MODEL EVALUATION

1. Introduction

This Appendix discusses the policies and procedures for the evaluation of nowcast/forecast models for navigation. The evaluation focuses on the performance of the model system during the development phases, and the accuracy of the system and its products during the operational phase. Since the suitability of the products is ultimately determined by the user (the navigational community), water levels and currents are the primary variables discussed here.

The main components of evaluation are standardization, periodic review, skill assessment, product quality control, and documentation. These are discussed in the next Section, but this Appendix will focus on skill assessment and the development of acceptance criteria (Section 3). Following that will be a discussion of the required performance measures for the evaluation for water levels (Section 4) and for currents (Section 5) when comparing predicted and observational data. However, if data are minimal or lacking altogether, or the product is an entire spatial field, other procedures are needed; these are discussed in Section 6. Once a system becomes semi-operational and nowcast/forecast products are being disseminated, additional procedures are necessary to assess the system; these product quality control procedures are discussed in Section 7.

2. Components of Evaluation

The main components of evaluation are Standardization, Periodic Review, Skill Assessment, Product Quality Control, and Documentation.

Standardization Standardization means that NOS model output and products, documentation, skill assessment, and review will be uniform. The standardization of products is covered in the Section 7, Product Quality Control. Although standardization is a goal, we must recognize that since each region and forecast system will be custom-tailored to some extent, we cannot expect total uniformity. However, each Developer should be aware of the features of other nowcast/forecast systems and have justification for significant departures.

Periodic Review During the model development, the Developer and the SDT will schedule and lead periodic (at least twice a year) technical meetings to present progress to the PFSSC. As part of the process, reviewers are to objectively assess progress and offer constructive written comments; the Developer is required to respond to written comments.

Skill Assessment Skill assessment is an objective measurement of how well the model nowcast or forecast does when compared to observations. The statistics were selected to measure the performance of the model in simulating astronomical tidal variability as well as total (tide and non-tidal effects) variability, to assess whether the model gives a more accurate forecast than the tide tables and/or persistence, and to gauge how well the model performs in an operational environment. Much of the skill assessment is completed before the nowcast/forecast products are disseminated. Skill assessment is discussed in depth in Section 3. If data are minimal or lacking, other procedures are needed (Section 6).

Product Quality Control Once the system begins operation, its products must be continually assessed for accuracy. To assist the Developer in judging the probable accuracy of the forecast, the forecast system will have standard products, save important information in a system status file, display the previous forecast, display the age of the oldest input data at the time of the forecast, and possibly display the probable forecast error (error bars). The last three items will help the user judge the probable accuracy. Details are discussed in Section 7.

Documentation Written documentation of the model and model system is necessary for the communication of model structure and performance. Documentation includes clear explanation of equations, boundary conditions, inputs, and constants, as well as sample (graphical) outputs and skill statistics. It will include copies of internally-documented computer code for all processes. It will also cover real-time and forecast input data streams and the telecommunications that provide them. See Appendix B for a further explanation.

3. Skill Assessment

Skill assessment is the primarily objective judgement of a model's performance (i.e., its ability to reproduce observed variability and to predict future variability) using a set of standards, and in relation to other prediction methods. The standards are set to meet user needs and to show how the prediction method can be improved, and are not influenced by the model's capabilities.

Skill assessment applies both to the model and to the entire nowcast/forecast system, since the availability, quality, and timeliness of input data (from observations and from other models) affects the quality of the nowcast/forecast. The methods discussed are to be applied to a model (either a numerical circulation model or statistical prediction scheme) that has been previously developed; therefore, basic questions about methodology (in the use of numerical models) about mass conservation, etc., will have been settled.

The components of skill assessment include: (1) the variables relevant to navigation, (2) the model run scenarios, (3) the comparison statistics or quantities, and (4) the acceptance criteria.

Relevant Variables In terms of importance to navigation, the primary variables are:

- the magnitude of the water level at all times and locations (for under-keel clearance),
- the times and amplitudes of high and low water at docking/anchorage sites,
- the speed and direction of the currents at all times and locations but especially at channel junctions (for maneuvering), and
- the start and end times of slack water before flood and ebb at all locations but especially at channel junctions (for planning turns in confined areas).

In terms of importance for scientific understanding and nowcast/forecast improvement, additional variables are the astronomical tides and tidal currents. Other potential variables (not addressed here) include water density (for cargo loading capacities) and maximum flood/ebb speed (for maneuvering). Skill assessment will be focused on the model system's accuracy in simulating the above variables.

The predicted and observed data must be processed to extract the variables important to navigation. In the comparison, the predicted data set is called P and the reference set is called R; P is produced by the model and R is produced from either the observed values or, for the tide-only scenario, from a tide prediction equation. The error, E, is P minus R. There are three Groups of P, R, and E:

- Group 1, a time series of values at uniform intervals
- Group 2, a set of values representing the consecutive occurrences of a variable/event,
- Group 3, a set of values representing a forecast valid at a given projection time

For Group 1, P1 and R1 will simply be the time series of values and E1 will be the difference. The series P1 and R1 may be filtered to subtract small-period variations. For currents, the time series will need to be in the following forms: (a) speed and direction and either (b) eastward and northward components or (c) along and counterclockwise normal to the principal current direction (PCD). The computation of the PCD requires eastward and northward components. In addition, some of the error statistics will be computed only under specific conditions. For example, the direction error is computed only for current speeds above ½ knot. For Group 2 (P2, R2, and E2), values are created from Group 1 series by selecting a sub-set of values such as the time and amplitude of high water and the time of the start and end of slack water (defined a having a current speed less than ½ knot) from P1 and R1. Group 3 (P3, R3,and E3) data are generated such that P3 consists of the values of the forecast variable that are valid at a fixed interval into the forecast (e.g., 0 hr, 6 hr, 12 hr, etc). The comparison series R3 is then the observed water level at the time the forecast is valid.

Model Run Scenarios The scenarios (Table B.1) explain the conditions under which the model is run, and are discussed in the order they would occur during model development. The scenarios begin with tidal simulations because tidal variations are generally dominant, they may account for a significant part of the error, and because there are extensive data available for validation. Modeled time series can be harmonically analyzed to produce constituent amplitudes and phases for comparison with accepted values. These values provide information on the model's behavior in frequency space and can also illuminate the role of friction and other non-linear processes.

The next scenario is the Test Nowcast, which is made with observed data with few (or filled in) data gaps. Similarly, the Test Forecast is made with input wind, boundary, and river forecasts that have few or filled-in gaps. Finally, the Semi-Operational Nowcasts and Forecasts are made in an operational environment (i.e., running daily) and so will occasionally encounter missing observations and forecasts; the system must be able to handle these conditions without significant loss of accuracy. During model development, the data streams used in the semi-operational runs should be saved so the Developer can test new prototype model formulations.

Skill Assessment Statistics Each Developer is responsible for generating a set of statistical values that will be used for model evaluation. Although no single set of statistics can quantify model performance perfectly, we have chosen from the available methods and selected easily-calculated quantities that provide relevant information on the important categories of model behavior. For a global assessment of errors, we use both the series mean (SM) and the frequency with which errors lie within specified limits (herein termed the central frequency, CF). The SM will indicate how well the model reproduces the observed mean and the CF indicates how often the error is within acceptable limits. The Root Mean Square Error (RMSE) and Standard Deviation (SD) are to be calculated, but have limited use since errors are not expected to be normally distributed and CF is easier to explain to users lacking a technical background. The frequency of times of poor performance is determined by analyzing the outliers, which are values which exceed specified limits. The Positive Outlier Frequency (POF) measures how often the nowcast/forecast is higher than the observed, a situation that should be minimized for water levels. The maximum duration of positive outliers (MDPO) indicates whether there are long periods when the model does poorly. The Negative Outlier Frequency (NOF) measures how often the nowcast/forecast is lower than the observed, a situation that should be minimized for currents at critical maneuvering locations. The maximum duration of negative outliers (MDNO) indicates whether there are long periods when the model does poorly. The MDPO and MDNO will be computed with data without gaps. The 'worst case', from a model-based nowcast/forecast viewpoint, is when actual water level turns out to be low but the model erroneously predicted much higher water levels and the user would have been better off using the astronomical tide water level prediction. A summary of relevant terms is shown in Table B.2.

Table B.1. Model Run Scenarios

Nowcast/forecast models are typically developed in the sequence explained below. In developing the Demonstration System, all scenarios will normally be run. In developing the Prototype System, the semi-operational mode will be run.

Scenario	Explanation
Tide Only	In this scenario, the model is forced with only harmonically-predicted astronomical tides for the open ocean boundary water levels. There are no wind stresses, river flows, or water density variations.
Test Nowcast	In this scenario, the model is forced with open ocean boundary water levels, wind stresses, river flows, and possibly water density variations that are based on observed, historical values. The observed values derive from nearly complete time series (i.e., minimal missing data).
Test Forecast	In this scenario, the model is forced with open ocean boundary water levels, wind, river flow, and water density variations based on forecast values from other models. The forecast values derive from nearly complete time series (minimal missing values or fields). Initial conditions are generated by observed data.
Semioperational Nowcast	In this scenario, the model is forced as in the Test Nowcast, but with actual observational input data streams. Significant portions of the data may be missing, so the model must be able to handle this.
Semioperational Forecast	In this scenario, the model is forced as in the Test Forecast, but with actual forecast input data streams. Significant portions of the data may be missing, so the model must be able to handle this.

Table B.2. Skill Assessment Variables.
Variables and statistics used in the skill assessment are explained below.

Variable	Explanation
Error	The error is defined as the predicted value p minus the observed (or reference) value r: $e_i = p_i - r_i$.
SM	Series Mean. The mean value of a time series of y. Calculated as $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$.
RMSE	Root Mean Square Error. Calculated as $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_i^2}$.
SD	Standard Deviation. Calculated as $SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N} (e_i - \bar{e})^2}$
CF(X)	Central Frequency. Fraction (percentage) of errors that lie within the limits $\pm$X.
POF(X)	Positive Outlier Frequency. Fraction (percentage) of errors that are greater than X.
NOF(X)	Negative Outlier Frequency. Fraction (percentage) of errors that are less than -X.
MDPO(X)	Maximum Duration of Positive Outliers. A positive outlier event is two or more consecutive occurrences of an error greater than X. MDPO is the length (number of consecutive occurrences) of the longest event.
MDNO(X)	Maximum Duration of Negative Outliers. A negative outlier event is two or more consecutive occurrences of an error less than -X. MDNO is the length (number of consecutive occurrences) of the longest event.
WOF(X)	Worst Case Outlier Frequency. Fraction (percentage) of errors that, given an error of magnitude exceeding X, that (1) the simulated value of water level is greater than the astronomical tide and the observed value is less than the astronomical tide or (2) the simulated value of water level is less than the astronomical tide and the observed value is greater than the astronomical tide.
PCD	Principal Current Direction. For an eastward current u and northward current v (Preisendorfer,1988), $PCD = \frac{1}{2}\arctan\left(\frac{2\sum_{i=1}^{N}(u_i-\bar{u})(v_i-\bar{v})}{\sum_{i=1}^{N}(u_i-\bar{u})^2 - \sum_{i=1}^{N}(v_i-\bar{v})^2}\right) + m\frac{\pi}{2}$ where m is either 0 or 1, whichever maximizes s^2 $s^2(PCD) = \cos^2(PCD)\sum_{n=1}^{N}(u_i-\bar{u})^2 + \sin(PCD)\cos(PCD)\sum_{n=1}^{N}(u_i-\bar{u})(v_i-\bar{v}) + \sin^2(PCD)\sum_{n=1}^{N}(v_i-\bar{v})^2$ PCD is counterclockwise from east and may represent either the flood or ebb direction.

Acceptance Criteria Many of the statistics described above have an associated acceptance criteria. For example, for water levels,

$$CF(15\text{ cm}) \leq 90\%,\ POF(30\text{ cm}) \leq 1\%,\text{ and } NOF(30\text{ cm}) \leq 1\%.$$

This means that 90% of the errors will be within $\pm$1/2 foot, only 1% of the errors will exceed 1 foot, and 1% of the errors will be less than 1 foot. This form of the criteria for the distribution of errors is used several times, and has the general form:

$$CF(X) \leq 90\%,\ POF(2X) \leq 1\%,\text{ and } NOF(2X) \leq 1\%.$$

The limit of 1% of the time is equivalent to about 87 hours per year (about 3½ days). Note that for a normal (Gaussian) distribution, the requirement that $CF(X) \leq 90\%$ implies that $SD = 0.608X$ and that $POF(2X)=0.05\%$. The general criteria for the duration (defined as the number of consecutive occurrences of an outlier) which can apply to data of Group 1, 2, or 3:

$$MDPO(2X) < N,\ MDNO(2X) \leq N$$

where N is the limit (N may or may not have units, depending on whether the data are from Group 1, 2, or 3). The above two sets of general criteria, plus some others, are required for the assessment of nearly all variables and are collectively called the Standard Suite (Table B.3).

For a nowcast or forecast at a particular station to be approved for release to the public, the statistics related to model performance at that station must (a) meet or exceed all criteria, or (b) meet or exceed most of the criteria and be granted a waiver by PFSSC. A waiver may be needed to allow for the wide variety of coastal areas and their dynamics, for changes in the priorities of users, and because a forecast is not likely to be as accurate as a nowcast. However, the basis for any waiver will be judicially considered. Legitimate bases include the fact that a time series of required length is unavailable and that a numerical criterion is missed by only a small fraction of the target value. The PFSSC may approve for dissemination a limited forecast (i.e., only a few forecast projections) or full (24 hour) forecasts for a limited number of locations. The PFSSC may choose to alter the accuracy requirements based on present day modeling capabilities. For example, the astronomical tide/current only simulation may be required to meet $CF(X) \leq 95\%$, the nowcast and 0-hr forecast may be required to meet $CF(X) \leq 90\%$, and the 6-, 12-, 18-, and 24-hr forecasts be required to meet $CF(X) \leq 85\%$, 80%,75%, and 70%, respectively. The outlier frequencies could be treated in a similar way.

The numerical values appearing in the criteria were selected on the basis of the estimated utility of the specific nowcasts/forecasts to users; it is expected that as nowcasts/forecasts become widely disseminated that some values will change. Statistics which have no specific acceptance requirement still are necessary for scientific model evaluation. To assist the decision-making process, all statistics generated from the data will be presented in a precision at least one place beyond that of the criterion or comparison value. The PFSSC will have final say on criteria and model acceptance.

Table B.3. Standard Suite and Standard Criteria for Skill Assessment
The table lists, for the appropriate data set, which skill parameters are required and which acceptance criteria (if any) apply. For the data set, P=predicted set, R=reference (observed or tide predicted) set, and E=error (P minus R). Definitions of skill parameters are as described in Table B.2.

Data Set	Standard Suite of Skill Assessment Parameters {Given X}	Standard Criteria for Acceptance {Given N}
P and R	SM	None
E	SM, RMSE, SD	None
	CF(X)	CF(X) $\geq$ 90%
	POF(2X), MDPO(2X)	POF(2X) $\leq$ 1%, MDPO(2X) $\leq$ N
	NOF(2X), MDNO(2X)	NOF(2X) $\leq$ 1%, MDNO(2X) $\leq$ N

The following sections list the evaluation criteria for water levels and currents, including the specific statistics and criteria, ordered by scenarios.

4. Evaluation Criteria for Water Levels

The following is a discussion of the statistics and acceptance criteria that are necessary for the evaluation of water levels. They are to be generated, analyzed, and documented by the Developer with a precision sufficient for an accurate comparison.

All time series of water level data shall consist of at least 365 days of hourly values; however, in cases where available time series of observations are less than 365 days, the PFSSC may make allowances. Minimum evaluation requirements for each scenario are as follows; they are summarized in Table B.4. In some locations, depending on navigational or other needs and on peculiar oceanographic requirements, additional requirements may be needed or some of the standard requirements may be relaxed.

Tide Only For the tide-only simulations, P1 is the model-based time series and R1 is the harmonically-predicted tide.

What is required for water levels is (1) a comparison of tidal constituents (37 amplitudes and phases[1]) harmonically analyzed from P1 with NOS accepted values and (2) the Standard Suite for X=15 cm and Standard Criteria for N=24 hrs. The X=15 cm requirement limits water level errors to within $\pm$1/2 foot and is based on our estimates of pilot's needs for under keel clearance. The SM for P1 and R1 are required so that the proper datum can be selected.

[1] The 37 constituents used by NOS are M_2, S_2, N_2, K_1, M_4, O_1, M_6, MK_3, S_4, MN_4, ν_2, S_6, μ_2, 2N, OO, λ_2, S_1, M_1, J_1, Mm, Ssa, Sa, Msf, Mf, ρ_1, Q_1, T_2, R_2, 2Q, P_1, 2SM, M_3, L_2, $2MK_3$, K_2, M_8, and MS_4.

Table B.4. Components of Model Skill Assessment for Water levels.
The table shows, for each model scenario and variable, which data sets are used, which skill assessment parameters must be calculated, and what acceptance criteria (if any) apply. Definitions of skill parameters are as described in Table B.1. WL is water level, THW is Time of High Water, TLW is Time of Low Water, AHW is amplitude at high water, ALW is amplitude at low water. P is the predicted set, R is the reference set, and E is the error (i.e., the difference between P and R). Subscript 1 refers to the uniform-interval time series, subscript 2 refers to a subset of 1 consisting of only a number of interpolated values taken from set 1, and subscript 3 refers to a subset of 1 that represents all values at a particular forecast valid time. BPA means that the values are better than the equivalent values for the persistence forecast and the astronomical tide-only forecast. For the Standard Suite and Standard Criteria, see Table B.3.

Variable	Set	Skill Assessment Parameters	Acceptance Criteria
		Scenario: Astronomical Tide Only	
WL	P1,R1	37 constituent amplitudes and epochs	None
	P1,R1,E1	Standard Suite {X=15cm}	Standard Criteria{N=24hr}
		WOF (30 cm)	WOF ≤ ½%
AHW, ALW	P2,R2,E2	Standard Suite {X=15cm}	Standard Criteria{N=3}
THW, TLW	E2	Standard Suite {X=30min}	Standard Criteria{N=3}
		Scenarios: Test Nowcast and Semi-operational Nowcast	
WL	P1,R1,E1	Standard Suite {X=15cm}	Standard Criteria{N=24hr}
		WOF(30cm)	WOF ≤ ½%
AHW, ALW	P2,R2,E2	Standard Suite {X=15cm}	Standard Criteria {N=3}
THW, TLW	E2	Standard Suite {X=30min}	Standard Criteria {N=3}
		Scenarios: Test Forecast and Semi-operational Forecast	
WL	E3	Standard Suite {X=15cm}	Standard {N=3}, BPA
		WOF(30cm)	WOF ≤ ½% , BPA
AHW, ALW	P2,R2,E2	Standard Suite {X=15cm}	Standard {N=3}, BPA
THW, TLW	E2	Standard Suite {X=30min}	Standard {N=3}, BPA

In addition, P1 and R1 are used to determine a second data set (P2 and R2) consisting of the amplitude of high water (AHW), the time of high water (THW), the amplitude of low water (ALW), and the time of low water (TLW); these are each accompanied by the error (E2). Precise times and amplitudes may be found by interpolation. Requirements on the amplitude data are (1) the Standard Suite for X=15 cm, Standard Criteria for N=3 hrs, and requirements on the times are (2) the Standard Suite (but without the SM for P and R) for X=30 min and Standard Criteria for N=3. The X=30 min requirement accounts for the fact that around the time of high (or low) water, the water level changes slowly.

Test Nowcast and Semi-operational Nowcast For these scenarios, P1 is the model-based time series (with all forcing included) and R1 is the observed time series (water levels must be referenced to the same datum). For the semi-operational nowcast, the individual nowcasts may be strung together to form a continuous series. For testing purposes, the semi-operational data streams with gaps must be preserved to recreate the characteristics of an operational environment. The MDPO and MDNO will be computed with data without gaps (but after missing data have been filled in).

For the total water level, what is required is the following: (1) the Standard Suite for X=15 cm and the Standard Criteria for N=24 hr. This should limit the consecutive occurrences to the approximate duration of a mid-latitude storm event. Also, (2) the worst case outliers, WOF(30cm) is required to meet WOF(30) $\leq$ 0.5%.

For the AHW and ALW data sets, what is required is (1) the Standard Suite for X=15 cm and the Standard Criteria for N=3. That is, there should be do more than three consecutive predicted high or low waters with an error greater than 30 cm. (Three consecutive semidiurnal high or low waters covers a period of about 24 hrs) For the THW and TLW data sets, what is required is (1) the Standard Suite for X=30 min and the Standard Criteria for N=3, but for the error only.

Test Forecast and Semi-operational Forecast Here there are series P2 (from each forecast) as before and P3, which is created by collecting all forecasts valid at a specific projection time (i.e., 0 hr, 6 hr, 12 hr, 18 hr, and 24 hr) from all 24-hr forecasts made each day. R is the observed value for the respective valid time.

For the test and semi-operational forecast scenarios, the statistics and variables needed for evaluation fall into two groups. The first group measures forecast skill using the Standard Suite and the Standard Criteria. The requirements for the water level error E3 are the Standard Suite for X=15 and the WOF(30 cm). The limit on the duration is N=3. For the series P2 and R2 of the amplitudes AHW and ALW, where P2 is determined from each forecast, the requirements are the Standard Suite for X=15 cm and the Standard Criteria for N=3. For the error E2 in the times THW and TLW, where P2 is determined from each forecast, the requirements are the Standard Suite for X=30 min and the Standard Criteria for N=3.

Although the acceptance requirements given here for the forecast are identical to those for the nowcast, it is recognized that in practice a forecast will generally be less accurate due to the uncertainties in the input. Potential relaxation of the acceptance requirements must rely on additional research.

The second group compares forecast skill using three methods: the model-based forecast, the astronomical tidal prediction with a persisted non-tidal component (i.e., based on an observed offset during some time before the forecast is made), and the astronomical tidal prediction alone. These are critical in deciding whether a model-based forecast should be issued at all. If the model forecast is not an improvement over the other two forecasts, the model should not be implemented. For testing purposes, the semi-operational data streams with gaps must be preserved to recreate the operational environment. The requirements are, for water level error and time of high/low water error, are

$$CF_{astronomical} \le CF_{persistence} \le CF_{model}$$
$$POF_{astronomical} \ge POF_{persistence} \ge POF_{model}$$
$$NOF_{astronomical} \ge NOF_{persistence} \ge NOF_{model}$$
$$MDNO_{astronomical} \ge MDNO_{persistence} \ge MDNO_{model}$$
$$MDPO_{astronomical} \ge MDPO_{persistence} \ge MDPO_{model,}$$
$$WOF_{astronomical} \ge WOF_{persistence} \ge WOF_{model}.$$

In other words, the model-based forecast gives better results than the persistence or the tide-only forecasts.

5. Evaluation Criteria for Currents

The following is a list of proposed statistics and acceptance criteria for the evaluation of currents They are to be generated, documented, and analyzed by the Developer with a precision sufficient for an accurate comparison.

The comparisons will be between the predicted time series (P) as compared to the reference time series, which is observed or harmonically-predicted (R) at specified depths at locations where observations exist. All time series of current meter data shall consist of 365 days of 6-minute (or 10-minute) data; however, in cases where available time series of observations are less than 365 days, the PFSSC may make an allowance. Currents will be at either prediction depth (15 ft below MLLW) or one-half the MLLW depth, whichever is smaller. Since the important variables are current speed, current direction, and the time of slack water, current time series will be needed in (1) speed and direction components and either (2) eastward and northward components or (3) along and perpendicular to the principal current direction (PCD).

Minimum evaluation requirements for each scenario are as follows; they are summarized in Table B.5. In some locations, depending on navigational or other needs and on peculiar oceanographic requirements, additional requirements may be needed.

Tide Only For the tide-only simulations, P is the model-based time series and R is the harmonically-predicted current.

For currents, what is required is: (1) for P1 and R1 having either eastward and northward components or components along and perpendicular the PCD, the SM and a comparison of tidal constituents (37 amplitudes and phases) harmonically analyzed from P1 with NOS accepted values; (2) for the current speed, the Standard Suite for X=26 cm/s (so that most errors will be less than ½ knot, based on pilot's needs for maneuvering) and Standard Criteria for N=24 hrs; (3) for current direction (provided the current speed is at least ½ knot), the Standard Suite for X=22.5 degrees (so that most errors will be less than half an octant of the compass) and Standard Criteria for N=24 hrs; and (4) for time of start and end of slack water (defined as a current speed not greater than ½ knot), the Standard Suite for X=15 minutes and Standard Criteria for N=3. For slack time errors E2, there will be two sets of the statistics, corresponding to slack before flood (SBF) and slack before ebb (SBE).

Table B.5. Components of Model Skill Assessment for Currents.

The table shows, for each model scenario and time series variable, which skill assessment parameters must be calculated, and what acceptance criteria (if any) apply. Definitions of parameters are as in Table B.1.SBF is slack before flood, SBE is slack before ebb. For the Standard Suite and Standard Criteria, see Table B.3.P is the predicted set, R is the reference set, and E is the error (i.e., the difference between P and R). Subscript 1 refers to the uniform-interval time series, subscript 2 refers to a subset of 1 consisting of only a number of interpolated values taken from set 1, and subscript 3 refers to a subset of 1 that represents all values at a particular forecast valid time. BPA means the values are better than the equivalent values for the persistence forecast and the astronomical tide-only forecast.

Series	Set	Skill Assessment Parameters	Acceptance Criteria
		Scenario: Astronomical tide	
Current	P1,R1	SM, 37 constituent amplitudes and epochs	None
Speed	P1,R1,E1	Standard Suite {X=26 cm/s}	Standard {N=24 hr}
Direction (speeds $\geq$ 26cm/s)	P1,R1,E1	Standard Suite {X=22.5 deg}	Standard {N=24 hr}
SBF, SBE (start time, end time)	E2	Standard Suite {X=15 min}	Standard {N=3}
		Scenarios: Test Nowcast and Semi-operational Nowcast	
Current	P1,R1	SM	None
Speed	E1	Standard Suite {X=26 cm/s}	Standard {N=24hr}
Direction (speeds $\geq$ 26cm/s)	E1	Standard Suite {X=22.5 deg}	Standard {N=24hr}
SBF, SBE (start time, end time)	E2	Standard Suite {X=15 min}	Standard {N=3}
		Scenarios: Test Forecast and Semi-operational Forecast	
Speed	E3	Standard Suite {X=26 cm/s}	Standard {N=3}, BPA
Direction (speeds $\geq$ 26cm/s)	E3	Standard Suite {X=22.5 deg}	Standard {N=3}, BPA
SBF, SBE (start time, end time)	E2	Standard Suite {X=15 min}	Standard {N=3}, BPA

Test Nowcast and Semi-operational Nowcast For these scenarios, P is the model-based time series (with all forcing included) and R is the observed time series. For the semi-operational nowcast, the individual nowcasts may be strung together to form a continuous series. For testing purposes, the semi-operational data streams with gaps must be preserved to recreate the characteristics of an operational environment. The MDPO and MDNO will be computed with data without gaps (but after missing data have been filled in).

For the total current, what is required is the SM for P1 and R1 (either eastward and northward components or components along and perpendicular to the PCD). There are no acceptance criteria for these.

For the current speed error E2, what is required is the Standard Suite for X=26 cm/s and the Standard Criteria for N=24 hours. This should limit the consecutive occurrences to the approximate duration of a mid-latitude storm event.

For the current direction, statistics are computed for times when the observed current speed is greater than 26 cm/s to get a meaningful measure of direction. P1 and R1 are time series consisting only of the values for the times when the speed criterion is met. What is required is the Standard Suite for X=22.5 degrees and the Standard Criteria for N=24 hours.

For the slack time, P2 is the set of predicted times of the start and end of slack water (defined as the current speed less than 26 cm/s) and R2 is the set of observed times for SBF and SBE. Precise slack times are computed by interpolation. What is required is the Standard Suite for X=15 min and the Standard Criteria for N=3. This requirement places most timing errors within a quarter of an hour (based on pilot's needs for planning a turn in a confined area) and limits the errors greater than a half an hour (30 min) to 1% of the time.

Test Forecast and Semi-operational Forecast Here there are a series P2 and several series P3, each created by collecting all forecasts valid at a specific projection time (i.e., 0 hr, 6 hr, 12 hr, 18 hr, and 24 hr) from all 24-hr forecasts made each day. R3 is the observed variable for the respective valid time. For these scenarios, the statistics and variables needed for evaluation fall into two groups.

The first group measures forecast skill just as for the nowcasts. For the speed error E3, the requirements is the Standard Suite for X=26 cm/s and the Standard Criteria for N=3. For the direction error E3 (provided speed is not less than 26 cm/s), the requirements is the Standard Suite for X=22.5 degrees and the Standard Criteria for N=3. For the slack time error, P2 and R2 are the first, second, third, etc. slack waters in the forecast series and in the observations for the same time period. The requirements are the Standard Suite for X=15 min and the Standard Criteria for N=3.

Although the acceptance requirements given here for the forecast are identical to those for the nowcast, it is recognized that in practice a forecast will generally be less accurate due to the uncertainties in the input. Potential relaxation of the acceptance requirements must rely on additional research.

The second group compares forecast skill using three methods: the model-based forecast, the astronomical tidal prediction with a persisted non-tidal component (i.e., based on an observed offset during some time before the forecast is made), and the astronomical tidal prediction alone. These are critical in deciding whether a model-based forecast should be issued at all. If the model forecast is not an improvement over the other two forecasts, the model should not be implemented. For testing purposes, the semi-operational data streams with gaps must be preserved to recreate the operational environment. The requirements are, for speed, direction, and slack water time errors,

$$CF_{astronomical} \leq CF_{persistence} \leq CF_{model}$$
$$POF_{astronomical} \geq POF_{persistence} \geq POF_{model}$$
$$NOF_{astronomical} \geq NOF_{persistence} \geq NOF_{model}$$
$$MDPO_{astronomical} \geq MDPO_{persistence} \geq MDPO_{model}$$
$$MDNO_{Atronomical} \geq MDNO_{persistence} \geq MDNO_{model}$$

6. Evaluation of Products at Sites Within an Estuary Which Lack Observational Data

User needs may require CO-OPS to issue nowcasts/forecast products for locations within an estuary where either real-time and/or historical observations are not available. Such products include both time series at individual locations and entire two-dimensional fields. The skill assessment of nowcast/forecast products at these locations must be handled differently than at locations where validation data are available. The procedures are as follows.

For a time series nowcast or forecast for a location where historical time series data exist, the model will be run to simulate the historical period (provided either wind data are available or wind effects are unimportant), or the historical data will be harmonically analyzed and compared to model-generated constituents. With constituent data only, model output will be harmonically analyzed for direct comparison.

For the case of locations where no historical data exist, the Design Team will analyze the model-generated data (time series or a field) and make a *professional judgement* as to its realism and the extent to which it captures actual features. The team will accomplish this by comparing the product to observational data from nearby stations and by assessing whether there are oceanographic reasons that it may be unrepresentative (e.g., location in an embayment separated from the main bay by a narrow, flow-restricting channel). This is especially important in assessing a current pattern which contains (or lacks) eddies and other features either known or hypothesized to exist.

The Design Team will present their results from the above analyses to the PFSSC. If the product is accepted, it can be disseminated for a trial period, but only with cautionary information. During initial dissemination, the nowcast field will be accompanied by a *cautionary note* (see below note 1) and will show the maximum of the errors at the locations with data for the time of the nowcast. A time loop will show the maximum error at the gauges for the period of the loop. A forecast field will be accompanied by a message showing/describing the forecast error bars. The Team will arrange for users to provide feedback on the accuracy and utility of the product. After a length of time sufficient for users to respond, the product will be reassessed by the Team for accuracy in light of comments. If revisions in the product or methodology are required, they will be implemented and the product will be issued for another trial period. If no revisions are necessary, the product will be disseminated with a *cautionary note* (see below note 2).

1. CAUTION: This product is presently being evaluated and should not be used for navigation or other official purposes. Probable error of water levels less than xx cm. (Probable error in current speed is less than xx cm/s and direction is less than xx degrees).

2. CAUTION: This product, although based on the best available information, should be used with caution for navigation or other official purposes. Probable error of water levels less than xx cm. (Probable error in current speed is less than xx cm/s and direction is less than xx degrees).

7. Product Quality Control

To assist the user in judging the probable accuracy of the forecast, the forecast system will have (1) standard products, (2) a system status file, and (3) a process to temporarily halt dissemination. It will display (4) the previous forecast, (5) the age of the oldest input data at the time of the forecast, and (6) the predicted forecast error (error bars).

Standard Products Each nowcast/forecast product will follow a set of standard specifications. The primary unit for water levels shall be feet and the datum will be Mean Lower Low Water. The primary unit for current speed will be knots and for current direction will be degrees clockwise from geographic north. Times will be local standard or daylight time (whichever is applicable) and have units of months, days, and hours on a 24-hour clock (e.g, Feb 1, 1300). The year will be shown. All maps will be in Mercator coordinates and will show latitude and longitudes.

System Status File This file will fulfill CORMS requirements and will catalogue the running time and completion status of all programs and the availability of input data fields (observed water levels, wind forecasts, and coastal water level forecasts). The file will show what data, if any, are missing and any other data or information that will indicate the quality of the nowcast and forecast for all stations. The file will also show the accuracy of the nowcast and previous forecast and keep a record of , for example, daily, weekly, monthly, and long-term model accuracy (see Tables B.3 and B.4) at all locations where data are available. Data in the file will be used to make a decision about whether to issue the nowcast/forecast product or to replace it with a message. These files will be saved for several days, or possibly archived, for assessment purposes.

Discontinued Dissemination CO-OPS needs to determine the conditions under which the nowcast or forecast remains valid. Using data in the system status file, CO-OPS will make a decision about whether to temporarily discontinue dissemination of some of the products. For example, the forecast may be witheld when no NWS wind forecasts are available. The decision to restore a temporarily discontinued product will be made by CO-OPS, in consultation with the SDT.

Previous Forecast The product (i.e., the nowcast for each location) will display the previous forecast values along with the observed values at the appropriate valid time so the user can judge the forecast accuracy. Model accuracy data will be saved in the System Status File.

Forecast Age The product will show the time at which the forecast was made and the age of the oldest input data. The age of the oldest input is selected from the ages of the meteorological forecast, the observed water level (for model initialization), and the coastal forecast.

Predicted Forecast Error Forecast products shall be accompanied by an indication of predicted forecast error for each forecast. The purpose of displaying the predicted error is to give the user a measure of the likelihood that the forecast will be reliable. Predicted error can be depicted graphically in two ways: as error bars or as an ensemble of forecasts. Error bars will show, at each forecast hour, the upper and lower limits that bound a fixed percentage (e.g., 90%) of all the test forecast results for that hour. *Event-dependent* error bars show the limits for a specific set of situations. For example, error bars may depend on whether the forecast winds over the region, when averaged over 12 hours, have a direction from either the north or the south. Other categories for developing event-dependent error bars could be based on a different averaging period, on averaged wind speed, or on the averaged non-tidal water level.

The ensemble approach requires several forecast runs, each with a different wind, ocean water level, or river flow forecast. For example, Eta-32 and Aviation model wind forecasts and either TDL or COFS water levels for the ocean forecast could be used. Each combination will produce a unique water level forecast. The upper and lower bounding forecasts would be the limits, with the mean at each hour forming the forecast. Note that each model and input combination would have to meet all skill assessment requirements.

The model Developer will generate each type of error bars and make a selection based on the reasonableness of the final product. Implementation of error bars will ultimately depend on the requirements of the user.